



UNITED STATES DEPARTMENT OF COMMERCE
Economics and Statistics Administration
U.S. Census Bureau
Washington, DC 20233-0001

The memorandum and attached document(s) was prepared for Census Bureau internal use. If you have any questions regarding the use or dissemination of the information, please contact the Stakeholder Relations Staff at dcco.stakeholder.relations.staff@census.gov.

2020 CENSUS PROGRAM INTERNAL MEMORANDUM SERIES: 2019.17.i

Date: May 1, 2019

MEMORANDUM FOR: The Record

From: Deborah M. Stempowski (signed May 1, 2019)
Chief, Decennial Census Management Division

Subject: 2020 Census Evaluation: Research on Hard to Count Populations: Non-English Speakers and Complex Household Residents including Undercount of Children Study Plan

Contact: Jennifer Reichert
Decennial Census Management Division
301-763-4298
jennifer.w.reichert@census.gov

This memorandum releases the final version of the 2020 Census Evaluation: Research on Hard to Count Populations: Non-English Speakers and Complex Household Residents including Undercount of Children Study Plan, which is part of the 2020 Census Program for Evaluations and Experiments (CPEX). For specific content related questions, you may also contact the authors:

Renee Ellis
Center for Behavioral
Science Methods
301-763-5122
renee.ellis@census.gov

Patricia Goerman
Center for Behavioral Science
Methods
301-763-1819
patricia.l.goerman@census.gov

Kathleen M. Kephart
Center for Behavioral Science
Methods
301-763-8891
kathleen.m.kephart@census.gov

Mikelyn V. Meyers
Center for Behavioral Science
Methods
301-763-9008
mikelyn.v.meyers@census.gov

Rodney Terry
Center for Behavioral
Science Methods
301-763-5475
rodney.terry@census.gov

Joanna Fane Lineback
Center for Behavioral Science
Methods
301-763-7599
joanna.fane.lineback@census.gov

Marcus P. Berger
Center for Behavioral Science
Methods
301-763-8335
marcus.p.berger@census.gov

United States Census 2020

2020 Census Evaluation

Research on Hard-to-Count Populations: Non-English
Speakers and Complex Household Residents Including
Undercount of Children

Study Plan

Renee Ellis, CBSM
Patricia Goerman, CBSM
Kathleen Kephart, CBSM
Mikelyn Meyers, CBSM
Rodney Terry, CBSM
Fane Lineback, CBSM
Marcus Berger, CBSM

Page intentionally left blank.

Table of Contents

I.	Introduction.....	1
II.	Background.....	2
III.	Assumptions.....	3
IV.	Research Questions.....	3
V.	Methodology.....	3
VI.	Subproject 1- Evaluation: NRFU Observations With Debriefing Interviews	5
VII.	Subproject 2- Bilingual Enumerator Training Experiment.....	15
VIII.	Data Requirements.....	21
IX.	Risks.....	21
X.	Limitations	22
XI.	Issues That Need to be Resolved	22
XII.	Division Responsibilities	23
XIII.	Milestone Schedule.....	24
XIV.	Review/Approval Table.....	24
XV.	Document Revision and Version Control History.....	25
XVI.	Glossary of Acronyms	25
XVII.	References.....	25

Page intentionally left blank.

I. Introduction

One particularly difficult aspect of the Census Bureau's mission is to include "hard-to-count" populations, including children, racial/ethnic minorities, and non-English speaking residents, in surveys and censuses. As the Census Bureau moves toward a person-based, records-based 2030 Census, we are confronted with the challenge of learning more about coverage in this context.

The purpose of this study is to conduct both an evaluation and an experimental research project about hard-to-count populations, such as non-English speakers and complex household residents during the 2020 census.

The evaluation research will aim to: (a) build on previous research about hard-to-count populations; (b) examine the interaction between the social context in 2020 and interviewer doorstep messaging to increase participation among hard-to-count populations; (c) conduct qualitative interviews to gain insights into issues relevant to non-English speakers and to develop a five minute doorstep debriefing to be conducted during the 2020 Census; (d) to observe doorstep interactions between interviewers and respondents across languages during the census and conduct debriefings to see what doorstep messages are most effective at increasing participation; (e) evaluate enumerator use of doorstep materials and hear from them about whether they have access to all that they need; (f) learn in real time about issues such as undercount of children, confidentiality concerns, and respondent propensity to respond across different survey modes.

For the evaluation research, in collaboration with a contractor, we will first conduct in-depth qualitative interviews in seven languages: English, Spanish, Chinese, Vietnamese, Korean, Russian and Arabic. The goal will be to gain in-depth insights into a variety of issues listed below, related to census participation and issues such as undercount of children and confidentiality concerns and to develop a five minute debriefing interview to take place in the field during observations of the 2020 Census enumeration.

For the experimental component of the research, we will develop an interviewer training module designed to support bilingual Spanish-speaking enumerators by providing them with research-based doorstep messages and practices to conduct interviews in Spanish. A test group of Spanish-speaking enumerators will receive the training for the 2020 Census, and outcomes such as response rates, contact attempts, and missing data will be compared across the test with a control group of bilingual interviewers who will not receive the training.

We will work with Field Division and the contractor with whom they are currently working to develop a brief, online Spanish bilingual interviewer training module that will have the same look and feel as the other 2020 Census training modules.

For both the evaluation and the experiment, we will select geographic areas with large non-English speaking populations where we will work with Field Division to match our contracted

bilingual observer ethnographers with bilingual enumerators during the census and to select interviewers for the experiment.

This proposal is meant to fill in gaps we have about coverage of hard-to-count populations and the reasons they are often missed during the decennial census. Many reasons may contribute to the undercoverage of these populations, but studies that take place outside of a decennial census year may fail to fully capture those reasons. The visibility of the decennial census creates a unique relationship between the government and its citizens that cannot be replicated in non-decennial years.

The information collected from this 2020 Census evaluation and experiment will inform research and innovations directed toward counting hard-to-count populations who typically have missing or limited administrative records available. The results will assist in planning throughout the decade leading up to 2030, so that we have support and methods in place to count these populations by the 2030 Census.

II. Background

The Census Bureau is interested in hard-to-count populations, including children, racial/ethnic minorities, and non-English speaking immigrants, where hard-to-count refers to known undercoverage of these populations in surveys and censuses. Multi method research at the Census Bureau has been used to study this topic with the goal of improving coverage for these populations. We know that for the 2020 Census, we can still expect many under-covered populations, including non-English speaking immigrants, to be less likely to complete the census via internet self-response than most of the population. The Non-Response Follow Up (NRFU) operation and face-to-face interviewing mode is extremely important for counting these populations.

As the Census Bureau moves towards a person-based, records-based 2030 Census, we are confronted with the challenge of learning more about the coverage issues of a mostly records-based census. For example, are we correct in our projections of who the “under-covered” populations are and where they will be in 2020? Do we understand their concerns and know the best ways to encourage their participation in the Census?

This proposed evaluation seeks to learn more about encouraging participation of populations that may be less willing or able to use online forms and those for whom these forms may pose more burden. Our primary focus will be on respondents who speak languages other than English. The areas in which these populations are concentrated will also include both English and non-English speakers who have historically higher rates of undercount of children in the Decennial Census. The goal is to gain insights both prior to and during the Census about these populations to help with understanding the 2020 census data and with planning the 2030 operations.

III. Assumptions

1. The project team will obtain adequate funding to implement the evaluation and experiment as it is designed in this study plan including all pretesting, travel, contractor support, and support for analysis and report writing.
2. The project team assumes that the Census Bureau will be able to obtain the services of a contractor to support the design and implementation of this evaluation/experiment.
3. CBSM will have access to paradata, response data and data from paper forms about the language the form was completed in and race/ethnicity of household respondents in order to analyze response rates and contact attempts for the experiment.
4. Training will be designed and developed to have a similar look and feel as the e-learning modules that enumerators will complete.
5. Geographic locations with similar demographic characteristics will be selected to enable comparisons between test and control groups for the enumerator training experiment.
6. FLD will recognize the selected Spanish bilingual enumerators and train them according to their appropriate treatment group.
7. Contractors and headquarters staff will observe live interviews during the 2020 Census to collect observations regarding the implementation of the training.
8. Enumerators will participate in debriefing focus groups following the 2020 Census.
9. Response data and paradata from the relevant geographic locations will be made available to research staff for statistical analysis of metrics such as contact attempts, response rates, item non-response, misreports, etc.
10. Response data will include an indicator for which enumerators worked on the case and whether they belonged to the treatment or control group.
11. Results and reports for external audiences will be reviewed by Data Products and Dissemination.

IV. Research Questions

Research questions vary by subproject and are detailed in the sub-project sections.

V. Methodology

A. Design

Methodology design varies by subproject and are detailed in the sub-project sections.

B. Interventions with the 2020 Census

This project has minimal impact on the 2020 Census; it does not require any interventions with existing solutions, systems or processes. There will be some interaction with field operations

staff and with enumerators, but these do not require any changes to existing systems or processes and we anticipate minimal impact.

The required activities for this project include:

1. Selection of the observation sites of NRFU interviews for the evaluation using the Planning Data Base (PDB), geographic information on language proficiency, and existing research on geographic areas likely to have large numbers of non-English speakers and undercounts of children. This selection process will use existing data sources and will not interfere in any way with Census processes
2. Sample selection of bilingual enumerators for the experiment into experimental and control groups using multistage sampling with existing census and PDB data. No intervention with any census processes will be required.
3. Field observations conducted in preselected areas. HQ staff and contractors will need Field's help to be assigned to NRFU interviewers for observations, including those who speak the target languages. Enumerators will be followed in the selected sites. The Center for Behavioral Science Methods (CBSM) understands that they will have no say in what households are selected for each enumerator. CBSM observers will accompany interviewers for whatever households are assigned on the given day. CBSM has met with FLD staff on several occasions to determine if they have any concerns with the impact of the observations or training experiment on operations during the 2020 Census. The FLD staff with whom we have spoken believe this research will minimally impact their operations and they were supportive of the projects. They have been given an opportunity to comment on the proposal at several points.
4. For the enumerator training experiment, collaboration with Field Division will be necessary in order to do a modification to the current Field training modules development contract for CBSM to develop the new training module in collaboration with the contractor. CBSM has spoken with Field training staff about this and they were supportive of the idea. CBSM will also collaborate with Field division to coordinate observation of enumerators in the test and control groups by contractor and headquarters staff. The training that enumerators receive in the treatment groups will not conflict with the standard training all enumerators receive. Rather, the additional modules will provide supplementary information to enumerators. The training will require approximately 30 additional minutes of training time for approximately 100 Spanish bilingual enumerators. Additionally, approximately 200 enumerators total will participate in 90-minute debriefing focus groups with 8-10 participants each. When CBSM consulted with Field Division staff, they encouraged CBSM to ensure that the training had the same "look and feel" as the training all enumerators receive in order to make the experience seamless, but otherwise felt that all parts were feasible.

C. *Implications for 2030 Census Design Decisions and Future Research and Testing*

As the 2030 Census moves toward an administrative records and self-response-based approach, hard-to-count households are particularly at risk of not being enumerated accurately via these approaches. Limited English Proficient (LEP) households in particular tend to fall into interviewer-administered modes, which has important cost implications for enumerating the U.S. population in 2030. This population is also likely to be missed with a records-based only 2030 Census and this study will provide insight into respondent concerns and strategies to increase response both at the doorstep and in other modes.

The expected outcome of our research would use recommendations and guidelines to improve undercounted population coverage, ethnic/racial minority coverage, coverage of young children. It also will provide an evaluation of 2020 privacy concerns that may be unique to the 2020 Census environment, such as response of hard-to-count populations to the citizenship question. This information can be used to inform content and operational decisions moving forward into 2030.

Other Expected Outcomes

- 1) Evaluated impact of the citizenship question on hard-to-count populations and research plans information for how to address messaging and strategies to increase response for future censuses and surveys that include this type of question.
- 2) Insights on rostering will show how well innovations in 2020 designed to improve rostering and reduce undercount worked and what additional areas may need improvement.
- 3) Insights into the use of paradata to identify respondents who are likely to use languages other than English to respond to the census.
- 4) Improvements in terms of coverage, data quality, and efficiency that may result from bilingual interviewers receiving specialized training could make enumerating LEP populations more accurate and cost effective in 2030. This training module could be adapted for languages other than Spanish and could be incorporated into other, non-decennial survey operations moving forward from these potential outcomes.

VI. Subproject 1- Evaluation: NRFU Observations with Debriefing Interviews

A. Background

We propose to conduct a 2020 Census evaluation of enumerator messaging and operations at the doorstep in the context of hard-to-count populations such as non-English speakers and members of racial and ethnic minority groups who have historically undercounted children at higher rates than the general population.

This evaluation will look at both our operations and at issues and concerns facing hard-to-count populations in the particular social context of 2020. This will allow us to examine how changes in procedure and process and new respondent concerns can be addressed going forward.

Subproject 1 will focus on hard-to-count populations of several types:

1. Limited English proficient (LEP) respondents (Martin and De la Puente 1993).
2. Complex household residents who may be monolingual English speakers, bilingual speakers of English and another language, or LEP respondents (Schwede and Terry 2013).
3. People who often omit babies, young children, and other household residents from their census forms.

There is a strong overlap in these three groups, and we therefore propose to study them together. By NRFU design, LEP households cannot be exclusively targeted, as such, some monolingual English households will end up in the study. Including questions on coverage allow us to take advantage of having an observer present and use this as a chance to study and ask meaningful questions of this population.

As a first step we will conduct qualitative research through cognitive interviews, both to gain insight and in-depth understanding of the point of view of people in these groups and to help us narrow down and prepare a list of no more than five minutes of doorstep, NRFU debriefing questions to ask in the field. Because we anticipate a strong overlap in residential areas of the groups of interest, we plan to come up with a sort of “tool kit” of debriefing questions that can be asked based on what is observed during NRFU interviews. For example, if the observer notices evidence of a child in the household, such as toys outside the home, but the respondent does not include children on the form, the observer can pull out the list of undercount questions to be included in the debriefing. If the observer notices a respondent refusing or unable to answer certain questions, he/she can pull out debriefing questions about sensitivity or difficulty with proxy response. Observation and debriefing at the time of the NRFU operation will provide a wealth of information about the current social context, interviewer preparedness, respondent concerns and information to help understand the data coming out of the 2020 Census and to help prepare plans and operations for 2030.

Previous 2020 Census evaluations on hard-to-count populations such as non-English speakers, complex household residents, and groups who have historically omitted young children have resulted in a lack of concrete and easily implemented action items. The report for this evaluation and experiment will include up-to-date information relevant to the current social context in 2020. In addition it will include specific messages, steps, and recommendations for each population studied and will include recommendations for products and processes that can be easily implemented in future operations. Finally it will include insights from qualitative findings that can be used to understand results of the 2020 Census and any anomalies that might arise, such as increased nonresponse or item nonresponse among specific populations. This information can also be used to tailor operations and outreach for the 2030 Census.

Examples of the types of insights this project will uncover are:

1. What challenges are enumerators facing in the current social context?

2. What kinds of paper materials do interviewers use and/or need at the doorstep? What are they missing?
3. How can respondents best be encouraged to include all young children on the census form at the doorstep?
4. What specific messages, actions, gestures of politeness are most successful in gaining cooperation of respondents who speak the languages in question?
5. What specific kinds of interviewer training materials are most useful? Which ones are currently lacking?

The decennial environment provides a unique opportunity to observe respondent privacy and confidentiality concerns given the mandatory nature of the census and the messaging campaign that explains the content of the census questionnaire. Respondents may be more aware of the presence of a citizenship question than they are in other survey instruments, and may have heard messages from community leaders regarding how they should respond to the requirement that they participate in the census. This qualitative research study has the potential to provide a wealth of information about these topics across cultural and language groups in the context of the 2020 Census.

Qualitative research may also provide insight into response patterns to the citizenship question in the production data of the census. While the census is different from many surveys because of its mandatory nature and because of the extensive outreach campaigns discussing the content of the survey instrument, qualitative findings on the topic of a citizenship question may also be somewhat applicable to other surveys containing citizenship questions.

1. Background Research on Respondents who Speak Languages other than English

During the 2010 Census, the Center for Survey Measurement (CSM), now called Center for Behavior Science Methods (CBSM),¹ conducted baseline NRFU observations for seven non-English languages. This research sought to collect qualitative data about sociolinguistic and cultural factors that affected enumeration, strategies used by enumerators to negotiate access to non-English speaking households, ways in which in-language census materials were used in the field, and how non-English speaking populations perceived the census and its public messaging.

Seven ethnographer teams went out to observe NRFU interviews, and they observed a total of 586 interviews (both in English and the non-English target languages). The ethnographers had previous experience conducting research in the communities and were fluent in the languages in question. The project was led by a team of Census Bureau coordinating researchers. Each multilingual team was then led by a contracted senior ethnographer and two to four assistant ethnographers. Ethnographers had backgrounds in a number of fields, including anthropology, sociology and linguistics. Ethnographers used an observation guide and debriefing protocol designed by the Census Bureau research team. Ethnographers were

¹ The Center for Survey Measurement (CSM) has been renamed to the Center for Behavior Science Methods (CBSM) as of October 2018. We use the new name for the remainder of this study plan even when referring to work carried out prior to the name change.

trained by the Census team and the team developed a comparative analysis framework to analyze results. There had been no official Census Bureau training for bilingual enumerators in terms of when and how to use non-English supplementary materials so an important part of the study was to see what enumerators were doing in the field.

In order to assist with 2020 Census research planning, the next section includes information about the number of in-language cases that were observed in the field in 2010 and how long it took to achieve that number of observations and debriefings. We include this information for five of the seven languages since individual reports were not published for the Vietnamese and Russian projects.

In 2010, observations for Arabic were carried out by three ethnographic observers in Michigan. Each observer was in the field for four to eight days. Each researcher observed up to 20 Arabic language cases, yielding a total of 60 observation/debriefings (Ajrouch et al. 2012).

Chinese NRFU interviews were observed by four ethnographers, who spent 19 days in the field in the metro Washington, D.C, area and in New York City. Of the 37 debriefings they conducted, 22 of them were in Chinese, mostly with Mandarin Chinese speakers, although a few spoke Cantonese (Shepherd et al. 2012).

Korean interviewers were followed by three ethnographers for 30 days in the metro Washington, D.C., area and in New York City. Of the total 83 NRFU interviews they observed, 23 were in Korean (Yoon et al. 2012).

The Portuguese observation and debriefings were carried out by three ethnographers who spoke very different dialects of Portuguese. These were conducted in southern New England over a period of 27 days. Of the 74 NRFU interviews observed, 33 were in Portuguese. These ethnographers reported a variety of language problems as they encountered very different regional dialects of Portuguese in their research site (Rodrigues et al. 2012).

Finally, Spanish observations were carried out by three ethnographers over 11 days in the field in Illinois. The observations focused primarily on Mexican and Mexican-American households in that region. Out of 88 NRFU interviews observed, 63 were Spanish-language cases (Isabelli et al. 2012).

On the whole, teams of three to four ethnographers in the field for two to four weeks in key areas with large language populations were able to observe 20 to 60 in-language cases during the NRFU field period.

General findings from these observations were: (1) cultural appropriateness and linguistic fluency of enumerators helped extend the length of interactions with non-English speaking respondents, (2) enumerators were more likely to go off script (deviating from the introductory messages included in the English instrument) when interviewing non-English speaking respondents than when interviewing English speakers, (3) enumerators did not have

knowledge about or access to available in-language materials prepared for the 2010 Census, (4) many interactions made use of on-the-fly translations or ad hoc interpreters, and (5) some non-English respondents were concerned that participation in the census would be prejudicial to them in some way.

In 2010 this research was conducted in seven non-English languages and different debriefing questions were asked of different communities, depending on the language being investigated, although the final reports were structured in the same way. This was due to a lack of coordination across the independent research teams and different ideas about what would be most relevant to ask about in each group. While this provided insight into tailored doorstep messaging for particular communities in 2020, it made direct comparisons between the languages' difficulty. Given the predicted changes in demographics by 2030, more general messaging and interviewer training for interacting with non-English respondents will be needed. In order to better evaluate interviewer procedures and issues that respondents are concerned about, we will be asking a standard set of debriefing questions across all languages including English. This will yield better comparison data across languages. Also, this type of research is important for directly understanding the barriers to completing NRFU interviews out in the field, especially for traditionally hard-to-count populations.

In 2016, CBSM conducted a large focus group study in which we filmed doorstep interaction examples using tailored messages to gain cooperation and participation in Spanish, Chinese, Korean, Vietnamese, Russian, Arabic and English. Respondents provided feedback on the messages and materials, such as the Language Identification Card, used to alert English only interviewers to the language spoken in a household. In addition to identifying particular messages that were most salient to members of each language group, one of the recommendations that came from this work was to create bilingual enumerator training and other supports to help achieve higher response rates for hard-to-count non-English speakers. CBSM has developed training for use in the 2018 Census test as well as for another current survey administered by the Census Bureau (the National Health Interview Survey or NHIS). Some of that training will be used in the training experiment detailed in subproject 2.

Findings from this research have been shared with various census stakeholders, and CBSM created a bilingual interviewer job aid that was given to a small subset of bilingual interviewers prior to the 2018 test. In addition, CBSM will be working with FLD in early 2019 to create two interviewer training modules for use in 2020: one on cross-cultural awareness and one for use with Spanish-speaking interviewers in Puerto Rico.

To address gaps in earlier studies and improve upon earlier work, the evaluation project will evaluate whether and to what extent non-English messages that have come out of various research projects over the last decade are getting into the field in 2020. Some examples of messages that can be used to encourage participation are: type of questions on the survey, interviewer legitimacy (showing Census ID), mandatory nature of the census, irrelevance of respondent immigration status to participation, confidentiality of the data, benefits to respondent communities, and reassurance about receipt of government benefits not being affected by survey participation.

By following NRFU interviewers in predetermined tracts likely to contain hard-to-count populations, this study will also seek to identify additional barriers and best practices for including non-English speakers in the count. Prior to observing the NRFU interviews, researchers will conduct a qualitative interviewing project designed to gain in depth knowledge about respondent views. These interviews will serve two additional purposes: 1. To pretest questions about a) respondent views on census materials, messages, modes, and barriers, and b) Undercount of children; and 2. To design and select five minutes worth of debriefing questions to be administered at the doorstep following NRFU observations.

As part of the debriefing, observers will ask no more than five minutes of additional questions after the interview (see Section B below for some preliminary plans on question topics). The additional questions will be carefully planned and pretested with cognitive testing on the target populations in early 2019. In 2020, tracts will be selected based on characteristics in the Planning Database (PDB) and other Census Bureau sources.

2. Background Research on Undercount of Children

Hard-to-count, limited English speakers are among the groups who often omit children from their census forms. We anticipate that we may see some examples of undercount of children through the course of our NRFU observation research. The respondent groups in question may also have useful insights about why children may be undercounted in their communities even if they do not leave any children off of their forms themselves. The reasons being, there will be little added cost to having questions on undercount ready to ask at the doorstep, and we can learn a lot about this topic in the qualitative interviews with this population prior to the field work. This study is a good opportunity to learn more about the undercount of children among hard-to-count respondent groups.

While the issue of undercounted children can and should also be studied outside of a decennial census environment, the unique visibility of a decennial census may yield important and unique insights. A decennial census is a highly publicized survey across languages and communities, which stresses the importance of including every person, yet historically census respondents have still left children age 4 and under off the form. Studying this issue is not the main purpose of this subproject, but we think that the project is a good opportunity to gain additional insights. These insights may be unique to the decennial census and interventions can be developed that will help future censuses.

Population estimates from demographic analysis methods (based on independent counts of births, deaths, and net migration) show that children under age 5 have been undercounted in the decennial census for decades (O'Hare 2015; U.S. Census Bureau 2014). The 2010 Census had an estimated net undercount of 4.6 percent young children, which amounts to about 1 million children under age 5 being missed (Konicki 2016; U.S. Census Bureau 2012).

Children under age 5 are at increased risk of omission if they live in large multigenerational or complex low-income households in disadvantaged communities (Fernandez et al. 2018;

O'Hare 2009; O'Hare et al. forthcoming; U.S. Census Bureau 2014). Researchers suggest that this is because shifts in demographic, social, and economic patterns have contributed to a separation between families and households, such that families may be spread over multiple households, or multiple families may be living in one household (Cherlin 2010). Thus, the increased complexity in household structures may cause ambiguity for census respondents about who to include on the household roster for a given residence, increasing the likelihood that some household members are not counted (U.S. Census Bureau 2016, U.S. Census Bureau 2017a, b, & c). Given the goal to move to an administrative records-based census in 2030, it is key that we understand more about these populations and ways to capture them in the census.

Other research (U.S. Census Bureau 2017d forthcoming) has also found that three specific types of complex households—households with nonrelatives, households with other unspecified relatives, and multigenerational households—accounted for 50 to 75 percent of undercounts within each of the single race (race-alone) and Hispanic origin groups. These results suggest that these types of households are at highest risk of undercounting and should be targeted in a future evaluation on the undercount of young children and other undercounted groups. Additional training has been added to this year's enumerator training about this topic. This includes information on rostering and a scenario about household rostering that is included in the practice training scenarios.

With limited federal resources, understanding and reducing the undercount of young children, as well as other under-covered populations, in the next census will require strategies such as identifying and targeting census tracts with a high concentration of people at risk of underenumeration. This research will help us understand rostering issues and other factors that impact the undercount of young children that can help us improve future survey work.

The undercount of young children and other household members is an issue that should be addressed in the 2020 Census, however, this is not the main goal of this subproject; the main goal is to study non-English speaking households. We know from past research that LEP populations are more likely to contain undercounted household members (Martin and de la Puente, 1993). In addition to the limited English speaking households in the areas we sample, we will encounter English only households in the selected sites. These households will be living in areas with lower cost housing and are likely to contain Hispanic, Black, and other non-White populations that past research has shown are more likely to contain extended family and/or unrelated members (Schwede, 2007). The results of this research can help inform decisions about whether or not to transition from a household-based to person-based collection in 2030 and what to do about populations that might be missed this way.

3. Background Research on Confidentiality and Privacy Concerns

A separate 2020 Census evaluation proposal focuses on using community partners to help understand confidentiality and privacy concerns of respondents who do not answer the census. However, with the inclusion of the citizenship question in the census, we expect to be able to learn about respondent concerns regarding this question as well as privacy

and confidentiality concerns more generally in this LEP evaluation project. We plan to discuss privacy and confidentiality concerns with respondents both in the qualitative interviewing phase of the project and at the doorsteps during the NRFU observation and debriefing study. During the NRFU operation, observers will track behaviors that might be indicative of respondent concerns with the citizenship question and other confidentiality concerns, including requests for clarification and refusals, as well as interviewer behavior during the administration of the citizenship question such as rephrasing the question or providing clarification. Additionally, observers will track any other privacy or confidentiality concerns raised by respondents and how interviewers react to these concerns.

B. Research Questions

The next section contains a list of all questions and topic areas that we propose to ask respondents in the qualitative interview/pretesting phase of the project. We will use that part of the project to gain in-depth insight into the topics and to develop and pretest five minutes worth of questions to ask respondents at the doorstep during the NRFU observation phase of the project.

Note: Topic area 1 below (language related questions) will be the main focus of observations in the field, but since it is also possible that we'll observe issues related to topics 2 (undercount of children) and 3 (confidentiality concerns) in the communities where we'll be observing, we plan to include these topics to see what insights we can gain. The qualitative interview study will help us choose five minutes worth of the most relevant debriefing questions. We plan to come up with multiple five minute lists of possible debriefing questions to choose from based on what the observer sees at each household. Debriefing questions will be prioritized according to the rarity of certain household characteristics to ensure that more unusual situations are addressed in debriefing before more common situations. We plan to conduct formal training with the observers to help them select the correct list of follow-up questions depending on the make-up of the household they are observing.

1. Language related questions
 - a. Observational questions:
 - i. Were bilingual enumerators able to convince respondents to participate at the doorstep?
 - ii. What messages did enumerators use to convince reluctant respondents? Was there variation with techniques and messages they used for English v. non-English respondents?
 - iii. What messages/techniques were not successful? Was there variation across the English and non-English cases they observed?
 - iv. What English-language or non-English written materials were shown to respondents at the doorstep?

- v. Was an interpreter needed/used? If so, who was the interpreter (household member vs. neighbor vs. professional interpreter)? What procedures were followed?
- vi. What difficulties occurred during the interview related to language barriers?
- b. Questions to pose to respondents:
 - i. How likely are Limited English Proficient (LEP) respondents to open mailings from the Census Bureau?
 - ii. Were any printed materials in their language clear/helpful?
 - iii. Would the respondent prefer to answer the census by mail, phone, or online (or other)? Why?
 - iv. Are there circumstances under which the respondent would feel comfortable completing the census online?
 - v. Was there any information that the respondent would like to have heard or received before participating in the census?

2. Undercount:

Note: Based on a 2010 Census evaluation, Schwede & Terry (2013) made recommendations to changing the under and overcount questions. Some of the recommendations were incorporated into the 2020 Census instrument. This is a chance to see how these changes worked. We would also like to examine the 2020 wording in the context of changing trends in household composition and differences in the political environment since 2010.

- a. Observational questions
 - i. Is there evidence of anyone else staying in the house that respondents left off the roster?
 - ii. Did respondents have any difficulty with rostering?
 - iii. Additional training was given on rostering to avoid undercounts. What techniques did enumerators use when completing the roster?
 - iv. Do the characteristics of people who are initially missed and later added match previous research findings?
 - v. Are there new demographic or cultural characteristics that haven't been observed before?
- b. Questions to pose to respondents
 - i. Are there any people who stay here that you were not sure if you should include? If there are, what could we have written or said that would have encouraged you to include this person or people?
 - ii. Are there young children like grandchildren or relatives who stay here sometimes or are staying here now? When would you count them as part of the household? When wouldn't you?
 - iii. Would someone else in the household responding include or exclude different people if they were answering?

3. Confidentiality concerns²: What privacy and confidentiality concerns do respondents have related to a decennial census environment?
 - a. Observational questions
 - i. What privacy or confidentiality concerns did respondents demonstrate, if any, related to participating in the census?
 - ii. Did respondents refuse to answer any census questions for themselves or other household members?
 - iii. Do respondent privacy or confidentiality concerns relate to participating in the 2020 Census vary by demographic group or household structure?
 - iv. How did enumerators address respondent privacy or confidentiality concerns?
 - v. Which enumerator strategies were most and least successful in addressing respondent privacy or confidentiality concerns relating to participating in the 2020 Census?
 - vi. How did respondents react to being asked about their citizenship status in particular?
 - vii. Did respondent reactions to being asked about citizenship status vary by demographic group or household structure?
 - viii. How did enumerators respond to concerns, if any, that respondents displayed about being asked their citizenship status?
 - ix. Which enumerator strategies, if any, were successful in overcoming respondent concerns about the citizenship question?
 - b. Questions to pose to respondents
 - i. Do you have any privacy or confidentiality concerns related to the 2020 Census? What are they?
 - ii. Do you know of other people who have specific privacy or confidentiality concerns?
 - iii. Did you have difficulty answering any of the census questions for yourself or others in your household?
 - iv. Does the citizenship question raise any concerns for you?
 - v. Have you heard others talking about the citizenship question in your community?

C. Methodology

There will be three components to this evaluation:

1. The first is qualitative research to delve into the research questions in depth and to develop a five-minute protocol for the later NRFU interviews. These qualitative interviewing sessions will be one-hour cognitive interviews conducted with the target populations. This part of the study will be a rich, qualitative data gathering exercise in its own right. It will also serve the secondary purpose of helping to plan and pretest the

² We plan to talk with people about this topic in the qualitative interview/pretesting phase. We will likely learn more during that phase of the project than we will at the doorstep given that those who are most concerned may not open the door at all. However, similar to the undercount of children topic, we would like to have some possible debriefing questions to pull out in the event that we do observe any evidence of these concerns at the doorstep.

NRFU debriefing questions. The interviews will be conducted in English, Spanish, Chinese, Vietnamese, Korean, Russian, and Arabic. This part of the project will occur in 2019 and will require contractor support. Results of the qualitative interviewing phase will be used to plan for the NRFU observation phase. For that phase we will develop an observation guide along with multiple five minute debriefings. Observers will be trained on how to use the observation guide and when and how to use each of the debriefing questionnaires. We will only use one minute questionnaire per household. Observer/researchers will be trained on how to select the best debriefing questionnaire for each type of situation they observe.

2. For the 2020 NRFU observation study, initially a potential 15 to 20 sites will be chosen based on meeting criteria of demographic characteristics associated with under-coverage and non-English speaking households. Sites will be chosen using the PDB and previous research on geographies that have high rates of undercount of young children. The criteria for potential sites will include percentage of households speaking each of the seven target languages, percentage of population that identifies as non-White, percentage of households in poverty, percentage of complex households, percentage of households with a child under 5 present, as well as other characteristics that previous research has shown are correlated with under-coverage. From the initial sites, five to seven sites will be strategically chosen. In the sites selected, we will work with field to identify three to five enumerators to pair with observers depending on the size of the site and availability. The target is approximately 25 pairings that will last five to seven days. Although we cannot guarantee how many interviews will actually occur and how many of those interviews will be in the target language, about 25 observations for each of the six languages (150 interviews) and at least 150 English interviews will be the goal.

Although we are using previous research and more quantitative methods to choose sites, these will not be random selections; they will be more strategic selections designed to increase our chances of observing interviews in households that have the characteristics we are interested in. As with all qualitative studies, we will seek as diverse a group as possible but will not make assertions of national representation.

3. During the 2020 Census, Census Bureau researchers and contractors will follow bilingual NRFU interviewers in these sites to observe interview interactions and conduct five-minute follow-up interviews with the target populations. Ideally, HQ staff and contractors will be able to audio record the entire NRFU interview, but at the very least, they can record the follow-up portion of the interview in addition to taking notes.

VII. Subproject 2- Bilingual Enumerator Training Experiment

A. Background

We propose conducting a 2020 Census experiment to evaluate the impact of specialized enumerator training on metrics like refusals, item non-response, and number of contact attempts in Spanish-language interviews during the Nonresponse Followup (NRFU) operation. This training will focus on enumerating Spanish-speaking households, and will be administered to a

group of Spanish-bilingual enumerators. In addition to statistical analysis, this mixed methods study will include qualitative field observations and enumerator debriefing focus groups.

Given the difficulty of capturing Limited English Proficient (LEP) households via administrative records or self-administered modes, streamlining field operations for enumerating LEP households will be an important component of accurately counting the U.S. population in the 2030 Census. This training could be adapted for other Census Bureau survey operations and into additional non-English languages.

The U.S. Census Bureau has devoted considerable resources over the past decade to developing and pretesting target-language materials and interviewer training with the goal of more accurately and efficiently enumerating LEP households, which are considered hard-to-count. In particular, CBSM has led the following research initiatives this decade:

- Ethnographic research in 2010 on enumerating Spanish-speaking households (Yuling Pan and Stephen Lubkemann, 2013)
- Debriefing focus groups with Census Questionnaire Assistance (CQA) enumerators from the 2015, 2016, 2017, and 2018 Census Tests ((Elizabeth Nichols et al, 2016; Erica Olmsted-Hawala et al, 2017; Elizabeth Nichols and Erica Olmsted-Hawala, 2018; Elizabeth Nichols et al, Forthcoming)
- Debriefing focus groups with Coverage Improvement (CI) and Re-Interview (RI) enumerators from the 2018 Census Test (Elizabeth Nichols and Erica Olmsted-Hawala, Forthcoming; Elizabeth Nichols et al, Forthcoming)
- Focus Groups Pretesting Enumerator Doorstep Introductions in 7 languages in 2017 (Anna Sandoval Girón et al, 2018)
- Focus groups with New York City Housing and Vacancy Survey (NYCHVS) interviewers in 2017 on training for administering the survey in non-English languages (Lucia Lykke et al, Forthcoming)
- Expert review of the Language Barrier enumerator training module for the 2018 Census Test (Mikelyn Meyers et al, 2018)
- Pretesting of messages relating to Census Participation in English and Spanish in 2018 (Aleia Clark Fobia et al, Forthcoming)
- Development of a bilingual enumerator handout on enumerating LEP households for usage in the 2018 Census Test (Lucia Lykke et al, 2018)
- Debriefing focus groups with 2018 Census Test NRFU enumerators on their experience with the bilingual enumerator handout developed by CSM
- Developing, administering, and revising a 1 hour Spanish bilingual training module for NHIS interviewers in preparation for their Redesigned NHIS Field Test in 2018 (Mikelyn Meyers et al, 2018; Mikelyn Meyers et al, 2018)
- Observation of Spanish bilingual NHIS interviewers during the Redesigned NHIS Field Test to determine the efficacy of the Spanish bilingual training they received (Mikelyn Meyers, Forthcoming)

- Focus groups with NHIS interviewers in 2017 and 2018 on training for administering the survey in Spanish (Mikelyn Meyers et al, Forthcoming; Mikelyn Meyers et al, Forthcoming)

Findings from the research conducted by Census Bureau staff listed above will inform the design of specialized training for our 2020 experiment, which represents the culmination of 10 years of research to provide additional support to bilingual interviewers. The experiment will be designed such that quantitative metrics will determine the extent to which the training interventions impacted response rates, data quality, and number of contact attempts in Spanish-language cases.

As the 2030 Census moves toward an administrative records and self-response-based approach, households that cannot be effectively enumerated via those approaches, i.e., those that fall into interviewer-administered modes, will increasingly drive the costs associated with enumerating the U.S. population. As such, efficiencies in terms of response rates and contact attempts as well as improving data quality for hard-to-count households will be increasingly important. If successful, this training could make Spanish bilingual enumerators more efficient at enumerating Spanish-language households. The training could be adapted for additional languages and used by other, non-decennial survey operations. Although research is driving the creation of training, that research is qualitative. A larger, more representative test will allow us to measure the benefit of this type of training before expending the money to give this training to all bilingual enumerators.

We have some limited qualitative data that indicates that the inclusion of the citizenship question may make some Spanish-speakers less willing to participate in the decennial census or to answer certain census questions due to their immigration status, or that of their household members (Aleia Clark Fobia et al, Forthcoming, CBSM, 2017). Although this hesitation may impact the response rates for these populations overall, we have no reason to believe response patterns would vary between the test and control groups in this experiment because of the procedure we will use to draw the random sample. We also consider this potential response pattern in our power analysis because it increases the complexity of our sample and will therefore mean we may need more cases in order to see significant differences.

B. Research Questions

1. Does Spanish training for bilingual enumerators impact response rates in Spanish-language interviews?
2. Does Spanish training for bilingual enumerators impact the number of contact attempts for Spanish-language interviews?
3. Does Spanish training for bilingual enumerators impact item non-response and misreporting errors in Spanish-language interviews?

4. Do observers see differences between enumerators with bilingual training and those without?
5. To what extent do Spanish bilingual enumerators in the test group report in focus groups that the additional training made them more or less successful at enumerating Spanish-speaking households? To what extent do Spanish bilingual enumerators in the control group report in focus groups that their training was or was not sufficient in preparing them to enumerate Spanish-speaking households?

We hypothesize that enumerator in the test group will be significantly more effective than enumerators in the control group (in regards to response rates, contact attempts, item nonresponse and reporting errors).

6. What do observers and focus groups report regarding the impact of the citizenship question? Does this differ between observations of control and test groups? We hypothesize that there will be no difference between the test and control groups in this regard but if we get qualitative evidence that there is a difference we can control for it in our analysis.

C. Methodology

Sample selection

In order to evaluate the research questions listed above, we propose an experimental design comparing a test and control group of enumerators in demographically similar locations.

We will use a multistage random sampling technique to select a sample of pairs of tracts with Limited English Proficiency (LEP) Spanish-speaking households using the Planning Database (PDB). We will use a matching algorithm to select tracts with similar demographic characteristics such as: Language, nativity, population density etc. In order to avoid overlapping test and control groups, we propose selecting two MSA's (Houston-The Woodlands-Sugar Land, TX, Metro Area and Dallas-Fort Worth-Arlington, TX, Metro Area). All control group tracts will be from 1 MSA and all test group tracts will be from the other. In order to counter the potential bias introduced by treatment groups across two separate MSA's, we will select a second sample of pairs from a third MSA (Los Angeles, CA). In this case, we will add an additional constraint to the matching algorithm to assure that test and control tracts are geographically distant from each other in order to avoid overlap. We did not include nativity or citizenship as demographic criteria because we believe LEP acts as a proxy for those characteristics. However, we will check our selected pairs to make sure they are similar in regard to those demographic characteristics and will add them as sample selection criteria if needed.

We will select all Spanish-bilingual enumerators in the selected pairs of tracts on the assumption that case assignment will attempt to assign cases close to the enumerator's home. Field Division has shared Spanish-bilingual recruitment goals with our team, with the caveat that the actual number of enumerators is expected to be fewer than the total number who are recruited, and that data from 2010 is not a good proxy for estimating the number of enumerators in 2020 because of changes in the recruitment procedures (e.g., recruitment is now an online process). As such, we

do not yet have a good estimate of how many bilingual enumerators will be in each geographic area, but we can reasonably assume that the selection of areas with high percentages of limited English proficiency Spanish-speaking households will yield a higher number of bilingual enumerators. We are estimating a sample of at least 40 pairs of tracts and we estimate at least 100 enumerators each in the test and control groups. Enumerators in the test group will receive the approximately 30-minute Spanish bilingual training. Enumerators in the control group will not receive the additional training module. The online training module will not be available to enumerators in the control group, which will prevent contamination between the groups. Census field supervisors working in the test and control sites will not be assigned this training, but will be briefed that some of their enumerators will be taking part in this experiment.

Power analysis

There are many unknowns that we have to contend with in this experimental design. For example, Field Division has a new method of adaptively assigning cases on a daily basis rather than assigning cases to a specific enumerator. Adaptive design introduces an increased potential for a particular case to be reassigned to one more other enumerators during the life cycle of the case. This is because cases are assigned daily rather than assigned to a particular interviewer. While we cannot model exactly how many enumerators will work the average case, we do our best to control for any effects by ensuring sufficient distance between test and control tracts to make it unlikely that the same case will be assigned to both an enumerator who received the training and one who did not during the life cycle of the case. Additionally, we cannot account for variability in supervisors, enumerator language skill or other enumerator characteristics. However, we do not have any reason to believe that these variations would be more likely to occur in the test or control group compared to the other group because these are both randomly sampled. There may also be underlying differences in test and control group populations that we may not know about and therefore can not account for in our sampling. However, a random sample should account for this variance as long as the sample size is large enough. In order to ensure that we have a large enough sample size to detect significant differences between the training and control groups, we designed a power analysis model to account for variation and guide how many tracts to select in order to have enough power. See attached paper detailing the power analysis model.

Training and evaluation

We will develop an approximately 30-minute online training module with the same look and feel as existing training that will cover the following topics:

- *Training for bilingual enumerators to improve data quality in non-English interviews*
 - Using the official translation rather than translating on the fly.
 - Bilingual materials available to enumerators.
 - Overcoming respondent concerns to avoid refusals and improve accuracy on questions that are frequently difficult for LEP respondents.

CBSM is currently assisting the Field Division and working with their training module contractor on two online training modules for use in the 2020 Census: 1) creating training to use for all enumerators on cross-cultural competence and 2) reviewing training for Spanish-speaking enumerators in Puerto Rico. We have spoken with the field training area about whether it would

be possible to do a modification on their current training module contract in order to create a special training module for use in this 2020 experiment. They have been very positive about this idea, which will help to ensure that the experimental module has the same look and feel as other training enumerators receive.

During the 2020 Census NRFU operations, researchers and contractors will follow enumerators in the field to observe Spanish-language cases completed by enumerators in both the test and control groups. We understand that enumerators are assigned cases via an adaptive design methodology and that Spanish-speaking enumerators may not encounter Spanish-speaking households on the dates of their observation. However, we are targeting geographic regions with high percentages of limited English proficiency Spanish-speaking households, which will maximize the likelihood that observers will encounter Spanish-language cases. Enumerators may be noncompliant with training instructions in the field, and indeed they are often noncompliant based on prior research. We have no reason to believe the test group will be more noncompliant than the control group. This is a limitation of conducting field research, but field observations will provide some insight into the extent to which enumerators implemented best practices covered in the training. Observers will track behaviors such as using the official translation rather than translating on the fly and enumerator success in overcoming respondent concerns. Observers will track behaviors such as using the official translation rather than translating on the fly and enumerator success in overcoming respondent concerns.

Toward the end of the 2020 Census NRFU operations, Spanish bilingual enumerators in the test and control groups will participate in debriefing focus groups to share the extent to which the training they received prepared them to enumerate Spanish-speaking households. We understand that as the field period progresses, enumerators may quit or be let go. We intend to coordinate these focus groups with field staff to ensure adequate numbers of enumerators in the test and control groups are still available to participate.

Analysis

Following the 2020 Census, CBSM will conduct statistical analysis of metrics between cases worked solely by enumerators in each group, as well as qualitative analysis of observational and focus group data. While cases worked by enumerators in more than one group will be excluded from analysis, we have created a sampling design that should minimize overlap between test and control groups. Cases completed with an observer present and cases completed by an enumerator who works as an interviewer on other Census Bureau surveys will be flagged and analyzed for their impact on the overall results. For example, career Census Bureau interviewers may have received specialized, Spanish-bilingual training during other survey operations that could contaminate results. We anticipate that excluding career Census Bureau interviewers from this experiment will have minimal impact on the study, as the majority of NRFU interviews will be completed by temporary interviewers hired specifically for the 2020 Census operation, given that career interviewers already have full caseloads on other survey operations. If these factors are found to effect the results, cases completed by these interviewers will also be removed.

Although we believe we have accounted for variations in geographic areas, demographics, enumerator and supervisor effects through our random sampling techniques, we will analyze the

data for any biases first. If these exist we can use modeling or weighting to adjust for biases before analysis.

Statistical analysis of the efficacy of the bilingual enumerator training will focus on response rates, number of contact attempts, response rates by contact attempt (to be able to distinguish between self reports and proxy reports), notes from interviewers about contact attempts, item nonresponse and item response selection on questions with high amounts of error for LEP populations.

Qualitative analysis from observational and focus group data will provide insight into what happened on the ground during the decennial census and will help explain quantitative trends in the data.

Once statistical and qualitative analysis are complete, CBSM will release a report documenting the efficacy of the specialized training in terms of effects on response rates, data quality, coverage, and cost and will provide a recommendation regarding the implementation of similar training in additional languages or survey operations.

VIII. Data Requirements

Data File/Report	Source	Purpose	Expected Delivery Date
PDB and other existing census sources	Publicly available data or already created data sets	Sample selection for subproject 1,2	2019
2020 Decennial response and paradata including enumerator id so we can flag enumerators in control and experimental group	2020 Decennial	Analysis for subproject 4	After collection period 2020

IX. Risks

1. If there is inadequate funding in FY 19, FY 20, or FY 21 then we will need to scale back evaluations and experiments.
2. If we are unable to partner with contractor(s) who can handle the seven target languages to conduct pretesting or the language NRFU follow-ups, then we will need to limit the number of languages we evaluate.
3. If we are not allowed to record the NRFU full interviews, then analysis will be limited to notes taken by observers.
4. If the hiring of Spanish-bilingual cases do not match projections, then we will have smaller experimental and control groups.

X. Limitations

Subproject 1:

1. Because of staff and resource limitations, we can't select a sample large enough to be nationally representative. The purpose of this research is to provide qualitative evidence only.
2. The NRFU followup interviews will need to be kept to a maximum of five minutes so questions will need to be concise and relevant. However, the qualitative/pretesting phase will give us more in depth insight into the issues.
3. We cannot be sure we will obtain successful NRFU interviews in any given day, or that the households will contain the target population.
4. This qualitative research on the effects of the citizenship question on respondent participation is not representative research and the findings will be limited in their generalizability to larger populations. However, since people who do not respond to the 2020 Census are not likely to be captured using other methods, this research will help to fill this gap.

Subproject 2:

5. If cases are worked by enumerators in both the test and control groups in the enumerator training experiment, they will have to be discarded because we can't determine if they should be in the control of experimental group. Cases observed by observers from headquarters or conducted by professional interviewers who work full-time on continuing surveys during non-census years would also be discarded if they are found to impact results. Both of these procedures may limit the pool of eligible cases for analysis, but we anticipate including enough cases in both treatment and control groups to overcome this limitation.
6. Control and experimental groups may be different in ways that were not detectable in the sampling design.

XI. Issues That Need to be Resolved

1. How OMB approval will be obtained to follow NRFU interviewers and ask additional questions?
2. How we coordinate with Field Division to identify interviewers to follow in the target sites.
3. CBSM needs to design and collaborate with NPC to access data centers training and data access processes?
4. Contracting and staffing- A contract needs to be put in place to get staff that speaks languages we are observing.
5. What mechanism will be used to assign a special training plan to enumerators in the test group?

XII. Division Responsibilities

Division or Office	Responsibilities
Field	<ul style="list-style-type: none"> • Helping to identify and pair observer/researchers with specific bilingual interviewers during the NRFU operation • Provide us with a copy of enumerator training and paper materials that enumerators are instructed to use with respondents in the field • Provide list of enumerators in test and control sites • Review training to ensure consistent look and feel • Ensure enumerators in test groups receive appropriate training • Coordinate focus groups
CBSM	<ul style="list-style-type: none"> • Planning, coordination and management of research project • Field work: conduct ethnographic observations and interviews • Contracting for additional field work • Select treatment and control sites for interview, select locations of observations and interviews • Design and program training • Analyze qualitative and quantitative data • Write final report with recommendations for 2030 and other Census Bureau operations
Decennial experiments and evaluations	<ul style="list-style-type: none"> • Coordinate and submit OMB package • Provide access to decennial data with appended information on interviewers who worked each case

XIII. Milestone Schedule

Milestones	Date
Develop supplementary training modules	04/2019-09-2019
Develop interview and observation draft protocols	4/2019-11/2019
Get contract in place for qualitative interviews to pretest observation protocols	9/2019-11/2019
Conduct Round 1 qualitative interviews and Round 2 qualitative interviews/pretesting of debriefing protocols for follow up interviews	11/2019-03/2020
Target geographic locations with similar demographic compositions in order to randomly select sites for test and control enumerators	9/2019
Select sites for NRFU observations	9/2019
Administer training to Spanish bilingual enumerators in the test group; Spanish bilingual enumerators in control group will not receive training	03/2020 – 04/2020
Observe enumerators in test and control groups	04/2020-06/2020
Observe NRFU interviews	
Conduct debriefing focus groups with enumerators in experiment test and control groups	06/2020-07/2020
Work on analysis of NRFU observations debriefing interviews	07/2020-02/2021
Receive, Verify, and Validate Data For language Experiment	09/2020-11/2020
Conduct statistical analysis of response metrics between cases worked solely by enumerators in test groups and cases worked solely by enumerators in the control group, as well as qualitative analysis of observational and focus group data	09/2020-05/2021
Distribute Initial Draft Report to the Decennial Research Objectives and Methods (DROM) Working Group for Pre-Briefing Review	06/2021
Decennial Census Communications Office (DCCO) Staff Formally Release the FINAL Report in the 2020 Memorandum Series	09/ 2021

XIV. Review/Approval Table

Role
Primary Author's Division Chief (or designee)
Decennial Census Management Division (DCMD) ADC for Nonresponse, Evaluations, and Experiments
Decennial Research Objectives and Methods (DROM) Working Group
Decennial Census Communications Office (DCCO)

XV. Document Revision and Version Control History

Version/Editor	Date	Revision Description
0.1/ RE	8/17/2018	First full draft
0.2/RE	8/20/2018	Full draft with comments from subproject leads
0.3/RE	9/26/2018	Draft with changes after feedback
1.0	9/26/2018	Submitted draft
1.1	1/15/19	Draft with comments from DROM
1.2	1/29/19	Draft with comments from leads
1.3	1/31/19	Full draft for DROM
2.0	3/14/2019	Revisions after process review
2.1	3/27/2019	Revisions after meetings with FLD
2.2	4/25/2019	Incorporate edit comments from DCCO

XVI. Glossary of Acronyms

Acronym
ADC
CBSM
DCCO
DROM
DSSD
EXC
IPT
R&M
FY
NRFU

XVII. References

Ajrouch, Kristine, Yuling Pan and Lubkemann. 2012. "Observing Census Enumeration of Non-English Speaking Households in the 2010 Census: Arabic Report." *Survey Methodology* 2012, 05.

Center for Survey Measurement. 2017. "Respondent Confidentiality Concerns." Memorandum for Associate Directorate for Research and Methodology. Sept 20, 2017.

Fernandez, Leticia, Rachel Shattuck and James Noon. 2018. The Use of Administrative Records and the American Community Survey to Study the Characteristics of Undercounted Young Children in the 2010 Census. CARRA working paper No 2018-05.

Fobia, Aleia Clark, Mikelyn Meyers, Aryn Hernandez, Lucia Lykke, (forthcoming) “Privacy Act Cognitive Testing High Level Findings and Recommendations,” CSM Research Report Series.

Hogan, H., Cantwell, P., Devine, J., Mule, V.T. and Velkoff, V. (2013) Quality and the 2010 Census. *Population Research and Policy Review*, Vol 32, No. page 637-662.

Isabelli, Christina, Yuling Pan and Lubkemann. 2012. “Observing Census Enumeration of Non-English Speaking Households in the 2010 Census: Spanish Report.” *Survey Methodology* 2012, 06.

Konicki, S. et al. (forthcoming 2017). Undercount of Young Children Research Team Report.

Lykke, Lucia, Gerson Morales, Mikelyn Meyers, Results from 2017 Focus Groups with Field Staff on Survey Translation and Limited English Proficiency (LEP) Households in the 2017 New York City Housing and Vacancy Survey (NYCHVS), CSM Research Report Series, Forthcoming.

Lykke, Lucia, Mikelyn Meyers, Patricia Goerman, “Guidelines for Counting Limited English-Speaking Households,” 2018, DX-697. Not publicly available.

Meyers, Mikelyn. “2018 Redesigned National Health Interview Survey Field Test CSM Observation Report,” CSM Research Report Series, Forthcoming.

Meyers, Mikelyn Cynthia Guerrero, Kim Geaghan, “2018 NHIS Bilingual Interviewer June Field Test Training,” 2018. Not publicly available.

Meyers, Mikelyn, Cynthia Guerrero, Kim Geaghan, Marcus Berger, “2018 NHIS Bilingual Interviewer September Training,” 2018. Not publicly available.

Meyers, Mikelyn, Kathleen Kephart, Jessica Holzberg, “2018 Redesigned National Health Interview Survey Focus Groups Report: Focus Groups with Monolingual and Bilingual Field Representatives Conducted During June 7 Training,” CSM Research Report Series, Forthcoming.

Meyers, Mikelyn, Lucia Lykke, Patricia Goerman, “Expert Review of NRFU E19 Language Barrier and Exercises Training Module,” 2018. Not publicly available.

Meyers, Mikelyn, Marcus Berger, Lidia Ojeda, “2018 Redesigned National Health Interview Survey Debriefing Report: Focus Groups with Monolingual and Bilingual Field Representatives Conducted on June 28, 2018,” CSM Research Report Series, Forthcoming.

Martin, Elizabeth and Manuel de la Puente (1993). “Research on Sources of Undercoverage Within Households.” Research Report series, 1993. Accessible at : <https://www.census.gov/srd/papers/pdf/sm93-03.pdf>, accessed on 12/7/18.

Nichols, Elizabeth, Erica Olmsted-Hawala, “Evaluating the Decennial Census Coverage Improvement Option: Results from an Interviewer Debriefing of the 2018 Census Test,” CSM Research Report Series, Forthcoming.

Nichols, Elizabeth, Erica Olmsted-Hawala, “Evaluating the decennial census call-in option: Results from an interviewer debriefing of the 2016 Census Test,” 2018, <https://www.census.gov/content/dam/Census/library/working-papers/2018/adrm/rsm2018-04.pdf>, accessed on 7/16/18.

Nichols, Elizabeth. Erica Olmsted-Hawala, Jonathan Katz, “Evaluating the decennial census call-in option: Results from an interviewer debriefing of the 2015 National Content Test,” 2018, <https://www.census.gov/content/dam/Census/library/working-papers/2016/adrm/rsm2016-02.pdf>, accessed on 7/16/18.

Nichols, Elizabeth, Erica Olmsted-Hawala, Mikelyn Meyers, “Evaluating the Decennial Census Reinterview Option: Results from an Interviewer Debriefing of the 2018 Census Test,” CSM Research Report Series, Forthcoming.

Nichols, Elizabeth, Erica Olmsted-Hawala, Mikelyn Meyers, “Evaluating the Decennial Census Call-In Option: Results from an Interviewer Debriefing of the 2018 Census Test,” CSM Research Report Series, Forthcoming.

O’Hare, William P. 2009. Why Are Young Children Missed So Often in the Census? Kids Count Working Paper. Baltimore, MD: The Annie E. Casey Foundation.

O’Hare, W. (2015). O’Hare W.P (2015). The Undercount of Young Children in the U.S. Decennial Census, Springer publishers.

O’Hare, William P., Deborah Griffin and Scott Konicki. (Forthcoming). Investigating the 2010 Undercount of Young Children – Summary of Recent Research. Report from the U.S. Census Bureau Task Force on the Undercount of Young Children.

Olmsted-Hawala Erica, Elizabeth Nichols, Mikelyn Meyers, CSM Suggestions for Updates to CQA based on 2017 Census test Interview Debriefings, Memorandum June 27, 2017. Not publicly available.

Pan, Yuling, Stephen Lubkemann, “Observing Census Enumeration of Non-English Speaking Households in the 2010 Census: Evaluation Report,” 2013, <https://www.census.gov/srd/papers/pdf/ssm2013-02.pdf>, accessed on 7/16/18.

Rodrigues, Isabel, Yuling Pan and Lubkemann. 2013. “Observing Census Enumeration of Non-English Speaking Households in the 2010 Census: Portuguese Report.” *Survey Methodology* 2013, 15.

Sha, M., Son, J., Pan, Y., Park, H., Schoua-Glusberg, A., Tasfaye, C., Sandoval Girón, A., García Trejo, A., Terry, R., Goerman, P., Meyers, M., and L. Lykke. 2018. “Multilingual Research for Interviewer Doorstep Messages Final Report.” Research Report Series. 2018. <https://www2.census.gov/library/working-papers/2018/adrm/rsm2018-08.pdf>, accessed on 7/16/18.

Schwede L. and R. Terry (2013). 2010 CPEX Evaluation. Comparative Ethnographic Studies of Enumeration Methods and Coverage across Race/Ethnic Groups in the 2010 Census. March, 2013.

Schwede L. (2007). “A New Focus Studying Linkages among Household Structure, Race/Ethnicity, and Geographic Levels, with Implications for Census Coverage.” Research Report Series. 2007, <https://www.census.gov/srd/papers/pdf/rsm2007-38.pdf>, accessed on 12/7/18.

Shepherd, Robert, Yuling Pan and Lubkemann. 2012. “Observing Census Enumeration of Non-English Speaking Households in the 2010 Census: Chinese Report.” *Survey Methodology* 2012, 08.

U.S. Census Bureau (2017a) Investigating the 2010 Undercount of Young Children – A Comparison of Demographic, Housing, and Household Characteristics of Children by Age, January 18, 2017.

U.S. Census Bureau (2017b) Investigating the 2010 Undercount of Young Children—Child Undercount Probes. January <date> 2017.

U.S. Census Bureau (2017c) Investigating the 2010 Undercount of Young Children – Examining Data Collected During Coverage Followup. January <date> 2017

U.S. Census Bureau (2017d) Investigating the 2010 Undercount of Young Children – Analysis of Complex Households.

U.S. Census Bureau (2017e) Investigating the 2010 Undercount of Young Children - Analysis of Coverage Followup Results Using the Esri Tapestry Segmentation and Planning Database. <insert date>.

U.S. Census Bureau (2017f) Call for 2020 Census Evaluation and Experiment Proposals. June 29, 2017.

U.S. Census Bureau (2016). Administrative Records, Internet, and Hard to Count Population Working Group of the National Advisory Committee on Racial, Ethnic, and Other Populations, U.S. Census Bureau, Washington DC.

U.S. Census Bureau. (2014). Final Task Force Report on the Undercount of Young Children, U.S. Census Bureau, Washington DC.

Yoon, Kyung-Eun, Yuling Pan and Lubkemann. 2012. "Observing Census Enumeration of Non-English Speaking Households in the 2010 Census: Korean Report." *Survey Methodology* 2012, 07.